

# THE TWO WEIGHT INEQUALITY FOR HILBERT TRANSFORM, CORONAS, AND ENERGY CONDITIONS

MICHAEL T. LACEY, ERIC T. SAWYER, CHUN-YEN SHEN, AND IGNACIO URIARTE-TUERO

**ABSTRACT.** We consider the two weight problem for the Hilbert transform, namely the question of finding real-variable characterization of those pair of weights for which the Hilbert transform acts boundedly on  $L^2$  of the weights. Such a characterization is known subject to certain side conditions. We give a new proof, simpler in many details, of the best such result. In addition, we analyze underlying assumptions in the proof, especially in terms of two alternate side conditions. A new characterization in the case of one doubling weight is given.

## CONTENTS

1. Introduction	1
2. Dyadic Grids and Haar Functions	4
3. Analysis of the Splitting Assumption	5
4. The Splitting of the Operator	9
5. The Calderón-Zygmund Corona	10
6. Bounded Fluctuation and the Second Corona	14
7. Dominating Functional Energy	18
8. The Remaining Estimates	20
References	23

## 1. INTRODUCTION

A *weight* is a non-negative Borel measure. We are interested in the two weight question for the Hilbert transform: For two weights  $(\sigma, \omega)$ , characterize the  $L^2$  inequality

$$(1.1) \quad \|H(\sigma f)\|_{L^2(\omega)} \leq B \|f\|_{L^2(\sigma)}.$$

Here, the inequality is understood in the sense that there is a uniform bound on the operator norm of a standard truncation on the singular integral kernel. Throughout, we will write  $H_\sigma f = H(\sigma f)$ , and understand at all times that some truncation is in place. The inequality above is in its self-dual

---

2000 *Mathematics Subject Classification.* Primary: 42B20 Secondary: 42B25, 42B35.

*Key words and phrases.* weights, Hilbert transform, corona decomposition.

Research supported in part by grant [nsf-dms 0968499](#).

Research supported in part by NSERC.

Research supported in part by the NSF, through grant DMS-0901524.

formulation: Interchange the roles of  $w$  and  $\sigma$  to get the dual inequality. We are also focused on  $L^2$  inequalities, so throughout we use the abbreviation  $\|f\|_\sigma := \|f\|_{L^2(\sigma)}$ . This conjecture, due to Nazarov-Treil-Volberg [12], has been the focus of attention.

**1.2. Conjecture.** *For a pair of weights  $(w, \sigma)$  we have the inequality (1.1) if and only if these three constants are finite.*

$$(1.3) \quad A_2 := \sup_{x \in \mathbb{R}} \sup_{t > 0} Pw(x, t) P\sigma(x, t),$$

$$(1.4) \quad H^2 := \sup_I \sigma(I)^{-1} \int_I |H(\sigma \mathbf{1}_I)|^2 w(dx),$$

$$(1.5) \quad H_*^2 := \sup_I \sigma(I)^{-1} \int_I |H(w \mathbf{1}_I)|^2 \sigma(dx),$$

where in the first line,  $Pw(x, t)$  denotes the Poisson extension of  $w$  to the upper half plane. In particular, the first line is an extension of the classical  $A_2$  condition, and is referred to herein as the  $A_2$  condition. The next two conditions are dual to one another, and are referred to as the testing conditions.

We will keep track of certain constants, like the three constants in the Conjecture above. Many of these will come in dual pairs, namely with the roles of  $w$  and  $\sigma$  reversed. An asterisk subscript will denote the dual constant, obtained through exchanging the roles of the two weights. The exact form of the Poisson integral is not important for us, and throughout we will use this form of it. For weight  $\sigma$  and interval  $I$ , we set

$$P(\sigma, I) := \int \frac{|I|}{(|I| + \text{dist}(x, I))^2} \sigma(dx).$$

This is the same, up to constants, as evaluating the usual Poisson extension of  $\sigma$  at the  $(c, |I|)$ , where  $c$  is any point of  $I$ .

To date, the Conjecture above has only been verified for pairs of weights which satisfy side conditions, which help control certain degeneracies in the weights  $\sigma$  and  $w$ . These side conditions are inspired by the *Pivotal Conditions* of [7], and were expanded and refined in [5], using the notion of *energy*. Our purposes are two-fold. (1) We will give a notably simpler proof of the best known current estimates. (2) We will analyze the proof strategy, introducing new side conditions sufficient for the two-weight estimate. These new side conditions are themselves, in a sense to be made precise in §3, a consequence of the correctness of the proof strategy. (3) We point out in Question 3.3, that it is not known if the proof strategy applies to all pairs of weights for which satisfy the two weight inequality. A new characterization when just one weight is doubling will follow from this analysis.

We define for interval  $I$  the *energy of  $w$  over  $I$*  to be

$$E(w, I)^2 := w(I)^{-2} \int_I \int_I \frac{|x - x'|^2}{|I|^2} w(dx) w(dx').$$

Assuming that  $|I| = 1$ , and  $w(I) = 1$ , this is two times the square of the distance, in the  $L^2(w \mathbf{1}_I)$  metric, of the function  $x \mathbf{1}_I$  to the linear space of constants. The *energy constant* of a pair of

weights  $(\sigma, w)$  is the smallest constant  $\mathbf{E}$  for which the following inequality holds. For all intervals  $I_0$  and all partitions  $\{I_j : j \geq 1\}$  of  $I_0$  we have

$$(1.6) \quad \sum_{j \geq 1} P(\sigma \cdot I_0, I_j)^2 E(w, I_j)^2 w(I_j) \leq \mathbf{E}^2 \sigma(I_0).$$

Here, inside the Poisson integral, we are identifying the interval  $I_0$  with its indicator function, which we will do throughout, as this will be a convenience in the heart of the proof.

A fundamental observation is that the energy constant is finite if the  $A_2$  constant and the testing conditions (1.4) and (1.5) hold. Namely, it was proved in [5] that we have  $\mathbf{E} \lesssim A_2 + \mathbf{H}$ . This depends upon the specific character of the  $1/y$  kernel; its modification for other relevant singular integrals is not nearly as simple.

We turn to the side conditions we need for our Theorem. Fix a choice of  $0 < \epsilon < \frac{1}{2}$ , and integer  $r \geq 2$ . We say that a pair of intervals  $(I, J)$  are  $(\epsilon, r)$ -good if for all  $J \subset I$ , satisfying  $|J| \leq 2^{-r}|I|$ , it follows that  $\text{dist}(J, \partial I) \geq |I|^{1-\epsilon}|J|^\epsilon$ .

**1.7. Definition.** The *Dini energy constant* of pair of weight  $(\sigma, w)$  is the smallest constant  $\Psi$  for which the following inequality holds: There is a decreasing non-negative sequence  $\psi(s)$  with  $\sum_{s \geq 1} \psi(s) = 1$ , so that for all integers  $s$

$$(1.8) \quad \psi(s)^{-2} \sum_{j,k \geq 1} P(\sigma \cdot (I_0 - I_j), I_{j,k})^2 E(w, I_{j,k})^2 w(I_{j,k}) \leq \Psi^2 \sigma(I_0).$$

In this inequality, we have these conditions.

- (1)  $I_0$  is an interval and  $\{I_j : j \geq 1\}$  a partition of  $I_0$ .
- (2) We have secondary partitions of  $I_j$  into intervals  $\{I_{j,k} : k \geq 1\}$ , so that the pair of intervals  $(I_j, I_{j,k})$  are  $(\epsilon, r)$ -good for all  $j, k \geq 1$ .
- (3) We have  $|I_{j,k}| < 2^{-s}|I_j|$  for all  $j, k \geq 1$ .

Note that here, it is certainly required that we consider the Poisson integral of  $\sigma$  restricted to the complement of  $I_j$ , else we could not expect to get the required decay in  $s$  to make the supremum finite.

This is very close to the side condition considered in [5], and is weaker than the pivotal condition of [7]. Namely, there is a pair of weights which *fail* one direction of the Pivotal Condition, but satisfy both directions of the side condition above, for  $\psi(s) \simeq 2^{-\epsilon s/2}$ , and the Hilbert transform is bounded for this pair of weights.

**1.9. Theorem.** Let  $w, \sigma$  be two weights which do not share any common point mass, and for some  $0 < \epsilon < 1$  and integer  $r$ , have finite Dini Energy Constant  $\Psi$ , and finite dual Dini Energy Constant  $\Psi_*$ . Then Conjecture 1.2 holds. Namely, we have the two weight inequality (1.1) if and only if the  $A_2$  condition (1.3), and the two testing conditions in (1.4) and (1.5) hold.

This theorem is essentially contained in [5], but the current proof contains many simplifications. Basic to the proofs are corona decompositions. We introduce herein a Calderón-Zygmund corona, whose use precludes the need for nuanced Carleson measure estimates. We still need a sophisticated corona decomposition modeled on one in [7], but in the current formulation we can

again avoid appeals to Carleson measure estimates. Prior arguments required a number of such arguments.

One of us initiated the study of two weight inequalities for the maximal function [10] and fractional integrals [11]. Cotlar and Sadosky have established two weight variants of the Helson–Szegő theorem [2], providing a complex analytic solution to the two weight problem. The dyadic variant of the Nazarov–Treil–Volberg conjecture is proved in [8]. Two weight inequalities for maximal truncations of singular integrals are studied by a completely different method in [4]. This paper represents, in some sense, a unification of these two lines of approach. The two weight problem for the Hilbert transform is closely related to a number of subjects, including embedding inequalities for model space [9] and de Branges space [1]; interpolating sequences for Paley–Weiner space [6]; and spectral theory for perturbed operators [3].

## 2. DYADIC GRIDS AND HAAR FUNCTIONS

**Dyadic Grids.** A collection of intervals  $\mathcal{G}$  is a *grid* if for all  $G, G' \in \mathcal{G}$ , we have  $G \cap G' \in \{\emptyset, G, G'\}$ . By a *dyadic grid* we mean a grid  $\mathcal{D}$  of intervals of  $\mathbb{R}$  such that for each interval  $I \in \mathcal{D}$ , the subcollection  $\{I' \in \mathcal{D} : |I'| = |I|\}$  partitions  $\mathbb{R}$ , aside from endpoints of the intervals. In addition, the left and right halves of  $I$ , denoted by  $I_{\pm}$ , are also in  $\mathcal{D}$ .

For  $I \in \mathcal{D}$ , the left and right halves  $I_{\pm}$  are referred to as the *children* of  $I$ . We denote by  $\pi_{\mathcal{D}}(I)$  the unique interval in  $\mathcal{D}$  having  $I$  as a child, and we refer to  $\pi_{\mathcal{D}}(I)$  as the  $\mathcal{D}$ -parent of  $I$ .

There is no unique choice of  $\mathcal{D}$ . To accomodate the notion of an interval being  $(\epsilon, r)$ -good, one must make a *random* selection of grids, but we have nothing to contribute to this portion of the proof. We refer the reader to [5, 7] for more details on this point.

**Haar Functions.** Let  $\sigma$  be a weight on  $\mathbb{R}$ , one that does not assign positive mass to any endpoint of a dyadic grid  $\mathcal{D}$ . We define the Haar functions associated to  $\sigma$  as follows.

$$h_I^{\sigma} := \sqrt{\frac{\sigma(I_-)\sigma(I_+)}{\sigma(I)}} \left( \frac{I_-}{\sigma(I_-)} - \frac{I_+}{\sigma(I_+)} \right).$$

In this definition, we are identifying an interval with its indicator function, and we will do so throughout the remainder of the paper. This is an  $L^2(\sigma)$ -normalized function, and has  $\sigma$ -integral zero. For any dyadic interval  $I^0$ , it holds that  $\{\sigma(I_0)^{-1/2}I_0\} \cup \{h_I^{\sigma} : I \in \mathcal{D}, I \subset I_0\}$  is an orthogonal basis for  $L^2(I_0\sigma)$ .

We will use the notation

$$\Delta_I^{\sigma} f = \langle f, h_I^{\sigma} \rangle_{\sigma} h_I^{\sigma} = I_+ \mathbb{E}_{I_+}^{\sigma} f + I_- \mathbb{E}_{I_-}^{\sigma} f - I \mathbb{E}_I^{\sigma} f.$$

The second equality is the familiar martingale difference equality, and so we will refer to  $\Delta_I^{\sigma} f$  as a martingale difference. It implies the familiar telescoping identity  $\mathbb{E}_J^{\sigma} f = \sum_{I: I \supset J} \mathbb{E}_J^{\sigma} \Delta_I^{\sigma} f$ . Finally, we will need the estimate below, which follows immediately from Cauchy–Schwartz.

$$(2.1) \quad |\mathbb{E}_{I_{\pm}}^{\sigma} h_I^{\sigma}| \leq \sigma(I_{\pm})^{-1/2}.$$

**Good-Bad Decomposition.** With a choice of dyadic grid  $\mathcal{D}$  understood, we then slightly change the definition of  $(\epsilon, r)$ -good. We say that  $J \in \mathcal{D}$  is  $(\epsilon, r)$ -good if and only if for all intervals  $I \in \mathcal{D}$  with  $|I| \geq 2^{r+1}|J|$ , we have that the distance from  $J$  to the boundary of *either child* of  $I$  is at least  $|J|^\epsilon |I|^{1-\epsilon}$ .

For  $f \in L^2(\sigma)$  we set  $P_{\text{good}}^\sigma f = \sum_{I \text{ is } (\epsilon, r)\text{-good}} \sum_{J \in \mathcal{D}^\sigma} \Delta_I^\sigma f$ . The projection  $P_{\text{good}}^w \phi$  is defined similarly.

Important elements of the suppressed construction of random grids [5, 7] are that

- (1) It suffices to consider a single dyadic grid  $\mathcal{D}$ , but we will sometimes write  $\mathcal{D}^\sigma$  and  $\mathcal{D}^w$  to emphasize the role of the two weights.
- (2) For any fixed  $0 < \epsilon < \frac{1}{2}$ , we can choose integer  $r$  sufficiently large so that it suffices to consider  $f$  such that  $f = P_{\text{good}}^\sigma f$ , and likewise for  $\phi \in L^2(w)$ .

Concerning the last property, this is, at some moments, an essential property. We suppress it in notation, however taking care to emphasize in the text those places in which we appeal to the property of being good.

### 3. ANALYSIS OF THE SPLITTING ASSUMPTION

Our principal concern is the bilinear form  $B(f, \phi) := \langle H_\sigma f, \phi \rangle_w$ ; let  $\mathbf{B}$  be the best constant in the inequality  $|B(f, \phi)| \leq \mathbf{B} \|f\|_\sigma \|\phi\|_w$ .

We define two more forms here. Throughout the paper by  $J \Subset I$  we mean that  $I, J$  are dyadic intervals, in a fixed dyadic grid, and  $J \subset I$  with  $|J| \leq 2^{-r}|I|$ , with  $r$  the fixed integer in the  $(\epsilon, r)$ -good property. Define

$$B_\epsilon(f, \phi) := \sum_{I \in \mathcal{D}^\sigma} \sum_{J \in \mathcal{D}^w : J \Subset I} \mathbb{E}_{I_J}^\sigma \Delta_I^\sigma f \cdot \langle H_\sigma I_J, \Delta_J^\sigma \phi \rangle_w$$

where  $I_J$  denotes the child of  $I$  that contains  $J$ . And, as mentioned in the previous section, we will identify an interval and its indicator function. Denote by  $B_\supset(f, \phi)$  the *dual* bilinear form obtained by interchanging  $w$  and  $\sigma$ . See Figure 1 for a diagram illustrating the definition of these two forms. Set  $\mathbf{B}_\epsilon$  be the best constant in the inequality

$$|B_\epsilon(f, \phi)| \leq \mathbf{B}_\epsilon \|f\|_\sigma \|\phi\|_w,$$

and  $\mathbf{B}_\supset$  be the best constant in inequality for the dual bilinear form.

In order to state our first main result, we need one more constant. Let  $\mathbf{W}$  be the best constant in the inequality

$$(3.1) \quad |\langle H_\sigma I, J \rangle_w| \leq \mathbf{W} \sigma(I)^{1/2} w(J)^{1/2},$$

where  $I$  and  $J$  are intervals with  $2^{-r}|J| \leq |I| \leq 2^r|J|$ . Recall that the integer  $r$  is fixed. It is known that  $\mathbf{W} \lesssim \mathbf{A}_2 + \min\{\mathbf{H}, \mathbf{H}_*\}$ .

**3.2. Theorem.** *Assume that the pair of weights satisfy the  $\mathbf{A}_2$  bound, and the two interval testing conditions (1.4) and (1.5). Then, assuming  $f = P_{\text{good}}^\sigma f$  and likewise for  $\phi$ , it holds that*

$$|B(f, \phi) - \{B_\epsilon(f, \phi) + B_\supset(f, \phi)\}| \lesssim \{\sqrt{\mathbf{A}_2} + \mathbf{W}\} \|f\|_\sigma \|\phi\|_w.$$

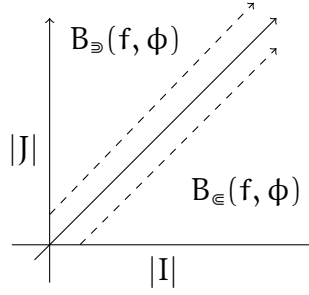


FIGURE 1. A schematic diagram for the two forms  $B_{\in}$  and  $B_{\sup}$ . The dashed lines around the diagonal indicate that the terms associated with  $2^{-r}|J| \leq |I| \leq 2^r|I|$  are treated in Theorem 3.2.

That is, the boundedness of  $H$  is equivalent to that of the sum  $B_{\in} + B_{\sup}$ . The remainder of the sufficiency proof for the main theorem is based upon the assumption that  $B_{\in}(f, g)$  and  $B_{\sup}(f, g)$  are *bounded independently of each other*. It is commonplace in classical settings that this assumption holds.<sup>1</sup> This brings up the following

**3.3. Question.** *Let  $w, \sigma$  be a pair of weights. Does it hold that  $B_{\sup} + B_{\in} \lesssim B$ ?*

Without an answer to this question, we cannot be sure that the approach to the two weight question used in this paper, and in [5, 7] can even succeed. Currently, there is no other approach to this question.

We introduce two new side conditions, more general, and more complicated, than the Dini condition; these conditions are phrased in terms of a dyadic grid, which is after all not fixed. The purpose in phrasing them is to provide precise description of those objects which the side conditions control.

**3.4. Notation.** For  $\mathcal{F}$  a subset of the dyadic grid  $\mathcal{D}$ , it is convenient to refer to  $\mathcal{F}$  as a sub-tree of the dyadic grid, and it useful to think of moving up or down the  $\mathcal{F}$ -tree, moving by inclusion. For dyadic  $I \in \mathcal{D}$ , we set  $\pi_{\mathcal{F}}I$ , the  $\mathcal{F}$ -parent of  $I$ , to be the minimal element of  $\mathcal{F}$  that contains  $I$ . We set  $\pi_{\mathcal{F}}^1 I = \pi_{\mathcal{F}}I$ , and inductively define  $\pi_{\mathcal{F}}^{t+1} I$  to be the minimal element of  $\mathcal{F}$  that strictly contains  $\pi_{\mathcal{F}}^t I$ . This has the consequence that if  $F \in \mathcal{F}$ , then  $\pi_{\mathcal{F}}^1 F = F$ . We write  $\text{Child}_{\mathcal{F}}(F)$  for the maximal elements of  $\mathcal{F}$  which are strictly contained in  $F$ , and call them the  $\mathcal{F}$ -children of  $F$ .

**3.5. Definition.** Given interval  $I^0$  we set  $\mathcal{F}(I_0)$  to be the maximal dyadic subintervals  $F$  such that  $\mathbb{B}_F^{\sigma}|f| > 4\mathbb{B}_{I_0}^{\sigma}|f|$ . We set  $\mathcal{F}_0 = \{I^0\}$ , and inductively set

$$\mathcal{F}_{j+1} := \bigcup_{F \in \mathcal{F}_j} \mathcal{F}(F).$$

Then, the collection of  $f$ -stopping intervals is  $\mathcal{F} := \bigcup_{j=0}^{\infty} \mathcal{F}_j$ .

<sup>1</sup>In various T1 theorems, there are canonical choices of paraproducts, which are bounded by the assumptions of the T1 theorem, whence they are freely added and subtracted in the proof. In the current setting, there is no canonical choice of paraproducts.

A basic fact, a consequence of the universal maximal function estimate, is

$$(3.6) \quad \sum_{F \in \mathcal{F}} \gamma(F)^2 \sigma(F) \lesssim \|f\|_\sigma^2, \quad \gamma(F) = \mathbb{E}_F^\sigma |f|.$$

This is referred to as the quasi-orthogonality condition.

We take  $f$  non-negative and supported on an interval  $I_0$ , and  $f$ -stopping intervals as above. Let  $\{g_F : F \in \mathcal{F}\}$  be a collection of functions in  $L^2(w)$  so that for each  $F$ ,

- (1)  $g_F$  is supported on  $F$  and constant on  $F' \in \text{Child}_{\mathcal{F}}(F)$ ;
- (2) letting  $\mathcal{J}^*(F)$  be the maximal intervals  $J^*$  such that  $J^* \subseteq F$ ,  $J^*$  is  $(\epsilon, r)$ -good, and  $\pi_{\mathcal{F}} J^* = F$ , we have  $\mathbb{E}_{J^*}^w g_F = 0$  for each  $J^* \in \mathcal{J}^*$ .

We say that  $g_F$  is  $\mathcal{F}$ -adapted to  $F$ . Let  $\mathbf{F}$  be the smallest constant in the inequality below, holding for all non-negative  $f \in L^2(\sigma)$ , and collections  $\{g_F\}$  as just described.

$$\sum_{F \in \mathcal{F}} \sum_{J^* \in \mathcal{J}^*(F)} P(f(\mathbb{R} - F)\sigma, J^*) \left\langle \frac{\chi}{|J^*|}, g_F J^* \right\rangle_w \leq \mathbf{F} \|f\|_\sigma \left[ \sum_{F \in \mathcal{F}} \|g_F\|_w^2 \right]^{1/2}.$$

We refer to this as the *functional energy condition*. Taking  $\mathcal{F}$  to be a partition of an interval  $I_0$ , and  $f = I_0$ , we can recover the energy condition (1.6). We denote by  $\mathbf{F}_*$  as the dual condition, with the roles of  $w$  and  $\sigma$  reversed.

The second condition is as follows. We write  $f \in \text{BF}_{\mathcal{F}}(F)$ , and say that  $f$  is of *bounded fluctuation* if (i)  $f$  is supported on  $F$ , (ii)  $f$  is constant on each  $F' \in \text{Child}_{\mathcal{F}}(F)$ , and (ii) for each dyadic interval  $I \subset F$ , which is *not contained in* some  $F' \in \text{Child}_{\mathcal{F}}(F)$ , we have  $\mathbb{E}_I^\sigma |f| \leq 1$ . We then denote as  $\mathbf{BF}$  the best constant in the inequality

$$(3.7) \quad \left| \sum_{I: \pi_{\mathcal{F}} I = F} \sum_{\substack{J: J \subseteq I \\ \pi_{\mathcal{F}} J = F}} \mathbb{E}_I^\sigma \Delta^\sigma f \cdot \langle H_\sigma I_J, \Delta_J^\sigma g \rangle_w \right| \leq \mathbf{BF} \{ \sigma(F)^{1/2} + \|f\|_\sigma \} \|g\|_w$$

where  $f \in \text{BF}_{\mathcal{F}}(F)$ , and  $g$  is  $\mathcal{F}$ -adapted to  $F$ . One must note that the two terms  $\sigma(F)^{1/2}$  and  $\|f\|_\sigma$  on the right above are in general incomparable. We refer to this as the *bounded fluctuation condition*.

This condition is a consequence of the boundedness of the form  $B_\epsilon$ , a fact which is not hard, and is proved below.<sup>2</sup> The role of the constant one in the inequalities  $\mathbb{E}_I^\sigma |f| \leq 1$  is immaterial. It can be replaced by any fixed constant. Indeed, if the measure  $\sigma$  is doubling, we could replace 1 by a constant depending only on the doubling constant, then the bounded fluctuation condition reduces to the function being in  $L^\infty$ .

The following Theorem summarizes much of the content of this paper.

**3.8. Theorem.** *The following inequalities and their duals hold, for any pair of weights  $w, \sigma$  which do not share a common point mass.*

$$(3.9) \quad \mathbf{B}_\epsilon \lesssim \mathbf{H} + \mathbf{F} + \mathbf{BF},$$

<sup>2</sup>But it is not known to us that the bounded fluctuation condition is a consequence of the norm boundedness of the Hilbert transform.

$$(3.10) \quad \mathbf{F} \lesssim \Psi, \quad \text{and} \quad \mathbf{BF} \lesssim \mathbf{H} + \Psi,$$

$$(3.11) \quad \mathbf{F} + \mathbf{BF} \lesssim \sqrt{\mathbf{A}_2} + \mathbf{W} + \mathbf{B}_\epsilon.$$

In particular, Theorem 1.9 is a corollary to the first two inequalities above, and their duals. The interest in (3.11) is that it shows that the new side conditions, of functional energy and bounded fluctuation, are implications of the proof strategy, namely the assumption that the bilinear form  $\mathbf{B}_\epsilon(f, \phi)$  is bounded. Concerning (3.10), the side condition controls the functional energy inequality by a straightforward argument, but the control of the bounded fluctuation term is a deep argument, §6, initiated in [7].

3.12. **Question.** *For a pair of weights  $(w, \sigma)$ , do any of these inequalities hold?*

$$\mathbf{F} \lesssim \mathbf{B},$$

$$\mathbf{F} \lesssim \sqrt{\mathbf{A}_2} + \mathbf{H},$$

$$\mathbf{BF} \lesssim \mathbf{B},$$

$$\mathbf{BF} \lesssim \sqrt{\mathbf{A}_2} + \mathbf{H}.$$

Note that the condition of functional energy is only about non-negative  $f$ , and the ‘energy’ of the weight  $w$ . It is arguably an acceptable hypothesis to add to Conjecture 1.2; unfortunately, neither functional energy nor bounded fluctuation conditions admit an intrinsic formulation. The inequality on bounded fluctuation goes to the heart of the conjecture.

Finally, we indicate a new characterization of the two weight problem when just one weight is doubling. This should be compared with the results of [4], which address maximal truncations, and also contrasts with a characterization in [7] when *both* weights are doubling.

3.13. **Theorem.** *Let  $(w, \sigma)$  be a pair of weights with  $\sigma$  doubling. Then, the two weight inequality (1.1) holds if and only if these constants are finite.*

$$\mathbf{A}_2, \mathbf{H}, \mathbf{H}_*, \mathbf{F}, \mathbf{BF} < \infty.$$

*Proof.* As  $\sigma$  is doubling, there is a constant  $c$  so that for any interval  $I$ , and any subinterval  $I'$  of length  $\frac{1}{4}$  of  $I$ , it holds that  $\sigma(I') \geq c\sigma(I)$ . From this, it follows that  $E(\sigma, I) \geq c/4$ . Namely, the energy of any interval is strictly bounded away from zero. Assuming the finiteness of  $\mathbf{A}_2, \mathbf{H}_*$ , as we may do in both directions of the argument, one may easily verify that the *dual* Dini condition holds, that is  $\Psi_* < \infty$ . (In fact, the pivotal condition of Nazarov-Treil-Volberg holds, as follows from the energy condition (1.6), which is necessary from  $\mathbf{A}_2$  and  $\mathbf{H}$ .)

Assuming that the Hilbert transform is bounded, we necessarily have the finiteness of the  $\mathbf{A}_2$  and testing constants. Therefore, the dual bilinear form is bounded,  $\mathbf{B}_\mathfrak{D} < \infty$ , hence  $\mathbf{B}_\epsilon$  is also finite, bounding  $\mathbf{F}$  and  $\mathbf{BF}$ , as claimed.

In the reverse direction, the finiteness of  $\mathbf{A}_2, \mathbf{H}_*$  and  $\Psi_*$  implies the boundedness of  $\mathbf{B}_\mathfrak{D}$ , and the additional assumptions on functional energy  $\mathbf{F}$  and bounded fluctuation  $\mathbf{BF}$  imply the boundedness of  $\mathbf{B}_\epsilon$ , hence the Hilbert transform is bounded.  $\square$



## 4. THE SPLITTING OF THE OPERATOR

We expand the full bilinear form  $B(f, g) := \langle H_\sigma f, \phi \rangle_w$  according to the weighted Haar basis. For the proof, we will take some (large) interval  $I^0$ , and assume that  $f$  and  $\phi$  are supported on  $I^0$ . Note that by the testing hypothesis,

$$|\mathbb{E}_{I^0}^\sigma f \langle H_\sigma(I^0), \phi \rangle_w| \leq \mathbf{H} |\mathbb{E}_{I^0}^\sigma f| \sigma(I^0)^{1/2} \|\phi\|_w.$$

The dual inequality also holds, so we are free to assume that  $f$  and  $\phi$  have respective means zero, and hence are in the closed linear span of the (good) Haar functions.

In the first generation, there are three terms, which are largely ‘below diagonal’, ‘diagonal’, and ‘above diagonal’ parts.

$$\begin{aligned} \langle H_\sigma f, \phi \rangle_w &= B_{1,1}(f, \phi) + B_{1,2}(f, \phi) + B_{1,3}(f, \phi), \\ B_{s,t}(f, \phi) &:= \sum_{(I,J) \in \mathcal{P}_{s,t}} \langle H_\sigma \Delta_I^\sigma f, \Delta_J^w \phi \rangle_w, \\ (4.1) \quad \mathcal{P}_{1,2} &:= \{(I, J) : 2^{-r}|I| \leq |J| \leq 2^r|I|\} \\ \mathcal{P}_{1,3} &:= \{(I, J) : |J| < |I|\}. \end{aligned}$$

The term  $B_{1,1}$  is dual to  $B_{1,3}$ , so we do not explicitly define it here, as we will concentrate on  $B_{1,3}$ .

The diagonal term is straightforward to control, and in §8.1, we will show

$$(4.2) \quad |B_{1,2}(f, \phi)| \lesssim (\mathbf{A}_2 + \mathbf{W}) \|f\|_\sigma \|\phi\|_w.$$

We shall follow this pattern of postponing certain estimates that are ‘routine’ till a later section, preferring to pass to the more delicate parts of the decomposition, which will always have the larger second indices.

We concern ourselves with the term  $B_{1,3}$  defined in (4.3). And we right away split it into an ‘far away’, ‘local’, and ‘inside’ part, defined as follows. Set

$$\begin{aligned} B_{1,3} &= B_{2,1} + B_{2,2} + B_{2,3} \\ (4.3) \quad \mathcal{P}_{2,1} &:= \{(I, J) \in \mathcal{P}_{1,3} : 3I \cap J = \emptyset\}, \end{aligned}$$

$$\begin{aligned} (4.4) \quad \mathcal{P}_{2,2} &:= \{(I, J) \in \mathcal{P}_{1,3} : J \subset 3I \setminus I\}, \\ \mathcal{P}_{2,3} &:= \{(I, J) \in \mathcal{P}_{1,3} : J \Subset I\}. \end{aligned}$$

In §8.2 and §8.3, we will show that these two terms are also controlled by the  $A_2$  constant.

$$(4.5) \quad |B_{2,1}(f, \phi)| + |B_{2,2}(f, \phi)| \lesssim A_2 \|f\|_\sigma \|\phi\|_w.$$

Concerning the term  $B_{2,3}$ , we will make this further decomposition. For the pairs of intervals  $(I, J)$  in question, we have  $J \subsetneq I$ . Recall that  $I_J$  is the child of  $I$  that contains  $J$ . Now, the argument of the Hilbert transform is  $\Delta_I^\sigma f$ , which is constant on the two children of  $I$ , namely  $I_J$  and  $I \setminus I_J$ . This permits us to write

$$B_{2,3} = B_{3,1} + B_{3,2},$$

$$(4.6) \quad \begin{aligned} B_{3,1}(f, \phi) &= \sum_{(I,J) \in \mathcal{P}_{2,3}} \mathbb{E}_{I \setminus I_J}^\sigma \Delta_I^\sigma f \cdot \langle H_\sigma(I - I_J), \Delta_J^w \phi \rangle_w, \\ B_{3,2}(f, \phi) &= \sum_{(I,J) \in \mathcal{P}_{2,3}} \mathbb{E}_{I_J}^\sigma \Delta_I^\sigma f \cdot \langle H_\sigma I_J, \Delta_J^w \phi \rangle_w. \end{aligned}$$

We will show in §8.4 that we have

$$(4.7) \quad |B_{3,1}(f, \phi)| \lesssim \mathbf{A}_2 \|f\|_\sigma \|\phi\|_w.$$

The bilinear form  $B_{3,2}$  is the form  $B_\epsilon$  of §3, and this is the notation that we will use below. Our considerations to this point, together with their duals, completes the proof of Theorem 3.2.

## 5. THE CALDERÓN-ZYGMUND CORONA

This section will be devoted to a proof of the inequality (3.9), namely that the side conditions of functional energy and bounded fluctuation can be used to control the bilinear form  $B_\epsilon$ . This is the first of the two important corona arguments in the paper. The reader should recall the definition of  $f$ -stopping intervals  $\mathcal{F}$  in Definition 3.5.

5.1. *Remark.* The intervals  $\mathcal{F}$  are the standard construct in proving paraproduct style arguments, moreover the identification and control of paraproducts is an essential part of the two-weight problem. Thus, it is natural to incorporate these intervals into the proof at an early stage. Indeed, if this step is not taken, nuanced Carleson measure estimates are needed.

5.2. **Definition.** [The Calderón-Zygmund Corona Decomposition] For  $F \in \mathcal{F}$ , we say that the pair of intervals  $(I, J) \in \mathcal{P}_{2,3}$  are in  $\mathcal{C}(F)$  if and only if  $\pi_{\mathcal{F}} J = F$ . This definition only depends upon  $J$ . We set  $\mathcal{C}_o(F)$  to be those pairs  $(I, J) \in \mathcal{C}(F)$  such that  $\pi_{\mathcal{F}} I_J = F$ . Note the dependence of this definition on the *pair*  $(I, J)$ . And, let  $\mathcal{C}^o(F) = \mathcal{C}(F) \setminus \mathcal{C}_o(F)$ . Define associated projections

$$P_F^w \phi := \sum_{J: \pi_{\mathcal{F}} J = F} \Delta_J^w \phi.$$

Note that the latter projections are pairwise  $L^2(w)$ -orthogonal in  $F \in \mathcal{F}$ , and we have

$$(5.3) \quad \sum_{F \in \mathcal{F}} \|P_F^w \phi\|_w^2 \leq \|\phi\|_w^2.$$

We use a similar, but distinct, notation  $P_F^\sigma f := \sum_{I: \pi_{\mathcal{F}} I_\pm = F} \Delta_I^\sigma f$ . Here, we sum over all  $I$  so that a dyadic child of  $I$  has  $\mathcal{F}$ -parent  $F$ . These projections are not orthogonal in  $F$ , but nevertheless satisfy a variant of (5.3) that we will need.

The (*Calderón-Zygmund*) *corona decomposition* of the bilinear form  $B_\epsilon$  is then based upon the  $f$ -stopping intervals, hence non-linear in nature.

$$B_\epsilon(f, \phi) := \sum_{F \in \mathcal{F}} \sum_{t=1}^3 B_t(f, \phi; F),$$

$$\begin{aligned}
(5.4) \quad B_1(f, \phi; F) &:= \sum_{(I, J) \in \mathcal{C}^0(F)} \mathbb{E}_F \Delta_I^\sigma f \cdot \langle H_\sigma(I_F \setminus F), \Delta_J^w \phi \rangle_w, \\
B_2(f, \phi; F) &:= \sum_{(I, J) \in \mathcal{C}^0(F)} \mathbb{E}_F \Delta_I^\sigma f \cdot \langle H_\sigma(F), \Delta_J^w \phi \rangle_w, \\
B_3(f, \phi; F) &:= \sum_{(I, J) \in \mathcal{C}_0(F)} \mathbb{E}_{I_J}^\sigma \Delta_I^\sigma f \cdot \langle H_\sigma(I_J), \Delta_J^w \phi \rangle_w.
\end{aligned}$$

Let us argue that we have equality above. The term  $B_3$  is the only one that depends upon  $\mathcal{C}_0$ , and it will be further decomposed below. The remaining two terms depend upon the complementary part of the corona  $\mathcal{C}$ . For  $(I, J) \in \mathcal{C}^0$ , note that  $I_J$  strictly contains  $F$ ,  $\Delta_I f$  is constant on  $I_J \supset F$ , hence  $\mathbb{E}_F \Delta_I^\sigma f = \mathbb{E}_{I_J} \Delta_I^\sigma f$ . And, we have written  $I_J = I_F = F + (I_F \setminus F)$  to get the two terms  $B_1$  and  $B_2$ .

**5.1. The Term  $B_3$ .** We will show in §6 that we have the inequality

$$(5.5) \quad |B_3(f, \phi; F)| \lesssim \{\mathbf{H} + \Psi\} \{\gamma(F) \sigma(F)^{1/2} + \|P_F^\sigma f\|_\sigma\} \|P_F^w \phi\|_w$$

An application of Cauchy-Schwartz, and (quasi)-orthogonality will complete the estimate of this term.

**5.2. The Term  $B_2$ .** We claim the estimate

$$(5.6) \quad |B_2(f, \phi; F)| \lesssim \mathbf{H} \gamma(F) \sigma(F)^{1/2} \|P_F^w \phi\|_w,$$

In view of the quasi-orthogonality condition (3.6) and (5.3), a trivial application of Cauchy-Schwartz will complete the estimate of this term. Namely, we have

$$\begin{aligned}
\sum_{F \in \mathcal{F}} |B_2(f, \phi; F)| &\lesssim \mathbf{H} \left[ \sum_{F \in \mathcal{F}} \gamma(F)^2 \sigma(F) \times \sum_{F \in \mathcal{F}} \|P_F^w \phi\|_w^2 \right]^{1/2} \\
&\lesssim \mathbf{H} \|f\|_\sigma \|\phi\|_w.
\end{aligned}$$

The proof of (5.6) is quickly obtained. We estimate, using the telescoping property of martingale differences,

$$\begin{aligned}
|B_2(f, \phi)| &\leq \sum_{J: \pi_{\mathcal{F}} J = F} \left| \sum_{I: (I, J) \in \mathcal{C}^0(F)} \mathbb{E}_F \Delta_I^\sigma f \right| \cdot |\langle H_\sigma F, \Delta_J^w \phi \rangle_w| \\
&= \left| \mathbb{E}_F^\sigma f \right| \|H_\sigma(F)\|_w \|P_F^w \phi\|_w \\
&\lesssim \mathbf{H} \gamma(F) \sigma(F)^{1/2} \|P_F^w \phi\|_w.
\end{aligned}$$

The expression  $\mathbb{E}_F \Delta_I^\sigma f$  arises above since for  $(I, J) \in \mathcal{C}^0(F)$ , it holds that  $F \subsetneq I_J$ . Hence, the sum of martingale differences can be summed exactly as above.

For  $B_3(f, \phi; F)$ , the definition of bounded fluctuation in (3.7) was constructed for this term. Namely, the function  $(C\gamma(F))^{-1} f \cdot F$  is in  $\mathbf{BF}_{\mathcal{F}}(F)$ . The function  $P_F^w \phi$  is  $\mathcal{F}$ -adapted to  $F$ , whence

$$|B_3(f, \phi; F)| \lesssim \mathbf{BF} \{\gamma(F) \sigma(F)^{1/2} + \|P_F^\sigma f\|_\sigma\} \|P_F^w \phi\|_w$$

An application of Cauchy-Schwartz, and appeal to (quasi-)orthogonality will complete the analysis of this term.

5.3. **The Term  $B_1$ .** The analysis of (5.4) is based upon the functional energy condition, and leads to this estimate:

$$(5.7) \quad |B_1(f, \phi)| \lesssim \mathbf{F} \|f\|_\sigma \|\phi\|_w.$$

The following lemma records a monotonicity property for the Hilbert transform, and a property involving the Poisson integral.

5.8. **Lemma (Monotonicity Property).** *Suppose that  $\nu$  is a signed measure, and  $\mu$  is a positive measure with  $\mu \geq |\nu|$ , both supported outside an interval  $I$ . Let  $J \subset J^* \Subset I$ . Then it holds that*

$$(5.9) \quad |\langle H\nu, h_J^w \rangle_w| \leq \langle H\mu, h_J^w \rangle_w$$

In addition, we have the estimate

$$(5.10) \quad P(\mu, J^*) \left| \left\langle \frac{x}{|J^*|}, h_J^w \right\rangle_w \right| \lesssim \langle H\mu, h_J^w \rangle_w + \left[ \frac{|J|}{|I|} \right]^{1-\epsilon} P(\mu, J) \sqrt{w(J)}$$

The function  $H\mu$  will be monotonically decreasing on  $J$ , and we have chosen the definition of the Haar functions so that  $\langle H\mu, h_J^w \rangle_w$  is non-negative, while  $\langle x, h_J^w \rangle_w$  is negative.

*Proof.* This argument is special to the Hilbert transform. Let  $J_- = J \cap (-\infty, c)$  and  $J_+ = J \cap (c, \infty)$ . We may renormalize the Haar function  $h_J^w$  so that

$$\int_{J_-} |h_J^w| dw = \int_{J_+} |h_J^w| dw = 1.$$

Then we have

$$\begin{aligned} \langle H\nu, h_J^w \rangle_w &= \int_{J_+} H\nu(x) h_J^w(x) dw(x) + \int_{J_-} H\nu(x) h_J^w(x) dw(x) \\ &= \int_{J_+} H\nu(x) |h_J^w(x)| dw(x) - \int_{J_-} H\nu(x') |h_J^w(x')| dw(x') \\ &= \int_{J_+} \int_{J_-} [H\nu(x) - H\nu(x')] |h_J^w(x')| dw(x') |h_J^w(x)| dw(x) \\ &= \int_{J_+} \int_{J_-} \int_{\mathbb{R} \setminus J} \frac{x - x'}{(y - x)(y - x')} d\nu(y) |h_J^w(x')| dw(x') |h_J^w(x)| dw(x), \end{aligned}$$

and since  $\frac{x - x'}{(y - x)(y - x')} \geq 0$  for  $y \in \mathbb{R} \setminus J$  and  $x \in J_+$  and  $x' \in J_-$ , we have

$$\begin{aligned} |\langle H\nu, h_J^w \rangle_w| &\leq \int_{J_+} \int_{J_-} \int_{\mathbb{R} \setminus J} \frac{x - x'}{(y - x)(y - x')} d\mu(y) |h_J^w(x')| dw(x') |h_J^w(x)| dw(x) \\ &= \langle H\mu, h_J^w \rangle_w, \end{aligned}$$

where the last equality follows from the previous display with  $\mu$  in place of  $\nu$ . This concludes the first half of (5.9).

For the second estimate (5.10), we will make a first order Taylor polynomial approximation of  $H\mu$  on the interval  $J$ . Let us denote the derivative by  $D$ , and for  $x \in J$  note that

$$DH\mu(x) = - \int \frac{1}{(x-y)^2} \mu(dy), \quad D^2H\mu(x) = \int \frac{1}{(x-y)^3} \mu(dy).$$

The point here is that the second derivative is somewhat small. From this, we can write, letting  $c_J$  be the center of  $J$ ,

$$\begin{aligned} |H\mu(x) - H\mu(c_J) - (x - c_J)H\mu(c_J)| &\leq \sup_{x \in J} (x - c_J)^2 |D^2H\mu(x)| \\ &\lesssim \frac{|J|}{\text{dist}(\partial J, I)} P(\mu, J), \quad x \in J. \end{aligned}$$

When we are estimating the inner product with a Haar function, constants are immaterial, therefore, using the fact that  $J$  is good, we can write

$$|\langle H\mu, h_J^w \rangle_w - DH\mu(c_J) \langle x, h_J^w \rangle_w| \lesssim \left[ \frac{|J|}{|I|} \right]^{1-\epsilon} P(\mu, J) \sqrt{w(J)}$$

Finally, one uses  $DH\mu(c_J) \lesssim |J|^{-1} P(\mu, J)$ . This finishes the argument.  $\square$

Returning to the analysis of  $B_1$ , write

$$\tilde{f}_F := \sum_{I: F \subseteq I} \mathbb{E}_{I_F} \Delta_I^\sigma f \cdot (I_F - F), \quad \bar{f} := \sum_{F \in \mathcal{F}} \gamma(F) \cdot F.$$

Note that  $|\tilde{f}_F| \lesssim \bar{f}$ . Let  $\mathcal{J}^*(F)$  be the maximal  $(\epsilon, r)$ -good intervals  $J^* \subseteq F$ . Applying the Lemma, we have

$$\begin{aligned} |B_1(f, \phi; F)| &\leq \sum_{J^* \in \mathcal{J}^*(F)} \sum_{\substack{J \subset J^* \\ \pi_{\mathcal{F}} J = F}} |\langle H_\sigma \tilde{f}_F, \Delta_J^w \phi \rangle_w| \\ &\lesssim \sum_{J^* \in \mathcal{J}^*(F)} \sum_{\substack{J \subset J^* \\ \pi_{\mathcal{F}} J = F}} P(\bar{f} \sigma(\mathbb{R} - F), J^*) \left| \left\langle \frac{x}{|J^*|}, \Delta_J^w \phi \right\rangle \right| \\ &= \sum_{J^* \in \mathcal{J}^*(F)} \sum_{\substack{J \subset J^* \\ \pi_{\mathcal{F}} J = F}} P(\bar{f} \cdot \sigma, J^*) \left\langle \frac{x}{|J^*|}, \Delta_J^w \bar{\phi} \right\rangle \end{aligned}$$

where  $\bar{\phi}$  is obtained from  $\phi$  by an appropriate  $w$ -Haar multiplier, chosen to make all the inner products above non-negative so that the absolute values can be removed. By orthogonality of the projections  $P_F^w$ , and the definition of the functional energy condition, we see that the sum over  $F \in \mathcal{F}$  of this last expression verifies (5.7).

To be specific, an operator  $T$  is a *w-Haar multiplier* if it is of the form  $T\phi = \sum_{J \in \mathcal{D}} \varepsilon_J \Delta_J^w \phi$ , with  $|\varepsilon_J| = 1$ . These operators are isometries on  $L^2(w)$ . The multiplier we need has  $\varepsilon_J = \text{sgn}(\langle \phi, h_J^w \rangle_w)$ , and  $\bar{\phi} = T\phi$ .

## 6. BOUNDED FLUCTUATION AND THE SECOND CORONA

There are two estimates of the bounded fluctuation constant  $\mathbf{BF}$  that should be made, the easy estimate of  $\mathbf{BF} \lesssim \mathbf{B}_\epsilon$ , and the difficult estimate of  $\mathbf{BF} \lesssim \sqrt{\mathbf{A}_2} + \Psi$ . We turn to the second estimate, which is involved.

Fix the data for the bounded fluctuation term.  $F$  is an interval, and  $\text{Child}(F)$  are the intervals inside  $F$ ; the function  $f$  is of bounded fluctuation relative to this data, and  $\phi$  is adapted to  $\{F\} \cup \text{Child}(F)$ . We consider the difficult estimate, in which the Dini and testing conditions dominate the bounded fluctuation term. Setting notation, we are to show (5.5), which is the same as this estimate.

$$\mathbf{B}_{\text{stop}}(f, \phi) := \sum_{I: \pi_{\mathcal{F}} I = F} \sum_{\substack{J: J \in \mathcal{I} \\ \pi_{\mathcal{F}} J = F}} \mathbb{E}_{I_J}^\sigma \Delta^\sigma f \cdot \langle H_\sigma I_J, \Delta_J^w \phi \rangle_w$$

$$(6.1) \quad |\mathbf{B}_{\text{stop}}(f, \phi)| \lesssim \{\mathbf{H} + \Psi\} \{\gamma(F) \sigma(F)^{1/2} + \|f\|_\sigma\} \|\phi\|_w.$$

The origins of this argument are derived from [7], as modified in [5]; these papers refer to this term as the *stopping term*. We will again find simplifications by the use of the Calderón-Zygmund corona. We define here the Dini corona.

**6.2. Definition.** Let  $I_0 \subset F$ . We set  $\mathcal{S}(I_0)$  to be the maximal subintervals  $S \subsetneq I_0$  such that

$$(6.3) \quad \Psi_w(I_0, S)^2 \geq 4\Psi^2 \sigma(S).$$

$$\Psi_w(I_0, S)^2 := \sup \psi(s)^{-2} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} P(\sigma \cdot (I_0 - I_j), I_{j,k})^2 E(w, I_{j,k})^2 w(I_{j,k})$$

where the supremum is formed over the various data that enter into Definition 1.7, to wit:

- $\{I_j : j \geq 1\}$  is a sub-partition of  $S$  into intervals;
- $\{I_{j,k} : k \geq 1\}$  is a sub-partition of  $I_j$  into good intervals,
- $s \geq r$  is an integer and  $|I_{j,k}| < 2^{-s}|I_j|$ , for all  $j, k$ .

We then set  $\mathcal{S} := \bigcup_{s=0}^{\infty} \mathcal{S}_s$ , where  $\mathcal{S}_0 = \{F\}$ , and inductively,  $\mathcal{S}_{s+1} = \bigcup_{S \in \mathcal{S}_s} \mathcal{S}(S)$ .

It is important to note that despite the assumption of the Dini energy condition, there is no *a priori* upper bound of the quantity  $\Psi_w(I_0, S)$  in terms of  $\sigma(S)$ . We also have the following elementary estimate, but critical,

$$(6.4) \quad \sum_{S \in \mathcal{S}(I_0)} \sigma(S) < \frac{1}{4} \sigma(I_0).$$

We have by (1.8),

$$4\Psi^2 \sum_{S \in \mathcal{S}} \sigma(S) \leq \sum_{S \in \mathcal{S}} \Psi_w(I_0, S)^2 \leq \Psi^2 \sigma(I_0).$$

The constant  $\Psi^2$  divides out, so that (6.4) holds.

The Dini corona decomposition of  $\mathcal{C}_0(F)$  is then the collection of pairs  $\mathcal{B}(S)$ , of those  $(I, J) \in \mathcal{C}_0(F)$  such that  $J$  has  $\mathcal{S}$ -parent  $S$ . We further write  $\mathcal{B}(S)$  as the disjoint union of  $\mathcal{B}_o(S) \dot{\cup} \mathcal{B}^o(S)$

where  $\mathcal{B}_0(S)$  consists of those pairs  $(I, J) \in \mathcal{B}(S)$ , where  $I_J \subsetneq S$ . This definition is carefully crafted so that (1.8) fails for  $I_J$  if  $(I, J)$  is in  $\mathcal{B}_0(S)$ .

We then split the term  $B_{\text{stop}}(f, \phi)$  up according to the corona. The argument of the Hilbert transform is also split up. Here, it is a basic fact that for each  $J$ , the function

$$(6.5) \quad b_J := \sum_{I: (I, J) \in \mathcal{B}_0(F) \cup \mathcal{B}^0(F)} \mathbb{E}_{I_J}^\sigma \Delta_I^\sigma f \cdot I_J$$

is supported on  $F$ , and has  $L^\infty$  norm dominated by 2. The Hilbert transform is applied to  $I_J$ . Let  $S$  be the  $\mathcal{S}$ -parent of  $J$ , We will write this as

$$I_J = \begin{cases} S - (S - I_J) & I_J \subsetneq S, \text{ equivalently, } (I, J) \in \mathcal{B}_0(S), \\ (I_J - S) + S & I_J \supset S, \text{ equivalently, } (I, J) \in \mathcal{B}^0(S). \end{cases}$$

And this permits us to write

$$(6.6) \quad \begin{aligned} B_{\text{stop}}(f, \phi) &= \sum_{S \in \mathcal{S}} B_1(f, \phi; S) + B_2(f, \phi; S) - B_3(f, \phi; S) \\ B_1(f, \phi; S) &:= \sum_{(I, J) \in \mathcal{B}(S)} \mathbb{E}_{I_J}^\sigma \Delta_I^\sigma f \cdot \langle H_\sigma(S), \Delta_J^w \phi \rangle_w \\ B_2(f, \phi; S) &:= \sum_{(I, J) \in \mathcal{B}^0(S)} \mathbb{E}_{I_J}^\sigma \Delta_I^\sigma f \cdot \langle H_\sigma(I_J - S), \Delta_J^w \phi \rangle_w \\ B_3(f, \phi; S) &:= \sum_{(I, J) \in \mathcal{B}_0(S)} \mathbb{E}_{I_J}^\sigma \Delta_I^\sigma f \cdot \langle H_\sigma(S - I_J), \Delta_J^w \phi \rangle_w \end{aligned}$$

**6.1. The control of  $B_3$ .** We take up the most delicate case of  $B_3(f, \phi; S)$ , showing that

$$(6.7) \quad |B_3(f, \phi; S)| \lesssim \Psi \|P_S^\sigma f\|_\sigma \|P_S^w \phi\|_w, \quad S \in \mathcal{S}.$$

Here, the projection  $P_S^\sigma f$  is onto the span of the Haar functions  $h_I^\sigma$  such that a child of  $I$  has  $\mathcal{S}$ -parent  $S$ , and  $P_S^w \phi$  has an analogous definition. Note that projections  $P_S^w$  are pairwise orthogonal, while a given Haar function can only contribute to at most two projections  $P_S^\sigma$ . This and application of Cauchy-Schwartz will show that

$$\sum_{S \in \mathcal{S}} |B_3(f, \phi; S)| \lesssim \Psi \|f\|_\sigma \|\phi\|_w,$$

which is as required in (6.1).

In the main estimate, we hold the relative lengths of  $I$  and  $J$  constant. It holds that

$$\begin{aligned} |B_{3,s}(f, \phi; S)| &:= \left| \sum_{\substack{(I, J) \in \mathcal{B}_0(S) \\ 2^s |J| = |I|}} \mathbb{E}_{I_J}^\sigma \Delta_I^\sigma f \langle H_\sigma(S - I_J), \Delta_J^w \phi \rangle_w \right| \\ &\lesssim \Psi \psi(s) \|P_S^\sigma f\|_\sigma \|P_S^w \phi\|_w, \quad s > r. \end{aligned}$$

The constants  $\Psi$  and  $\psi(s)$  are as in Definition 1.7, and in particular,  $\sum_s \psi(s) \leq 1$ . This is summed over  $s$  to finish the proof of (6.7).

To prove the inequality above, we use this observation. For any choice of sign,

$$|\mathbb{E}_{I_{\pm}}^{\sigma} \Delta_I^{\sigma} f| = |\langle f, h_I^{\sigma} \rangle_{\sigma}| |\mathbb{E}_{I_{\pm}}^{\sigma} h_I^{\sigma}| \leq |\langle f, h_I^{\sigma} \rangle_{\sigma}| \sigma(I_{\pm})^{-1/2}.$$

This is the elementary property of the Haar functions of (2.1). We apply Cauchy-Schwartz in  $I$  to the expression below

$$\begin{aligned} |B_{3,s}(f, \phi; S)|^2 &\lesssim \|P_S^{\sigma} f\|_{\sigma}^2 \times \sum_{I: \pi_S I = S} \sigma(I_J)^{-1} \left[ \sum_{J: \substack{(I,J) \in \mathcal{B}_1(S) \\ 2^s |J| = |I|}} |\langle H_{\sigma}(S - I_J), \Delta_J^w \phi \rangle_w| \right]^2 \\ &\lesssim \|P_S^{\sigma} f\|_{\sigma}^2 \cdot \|P_S^w \phi\|_w^2 \times M_s^2 \\ M_s^2 &:= \sup_{I: \pi_S I = S} \sup_{\theta \in \{\pm\}} \sigma(I_{\theta})^{-1} \sum_{\substack{(I,J) \in \mathcal{B}_0(S) \\ J: \substack{2^s |J| = |I| \\ I_J = I_{\theta}}}} \langle H_{\sigma}(S - I_{\theta}), h_J^w \rangle_w^2. \end{aligned}$$

Here, to get the bound in terms of  $\|P_F^w \phi\|_w^2$ , we use the fact that for fixed  $J \subset S$ , there is a unique  $I$  so that  $J \subset I$ ,  $2^s |J| = |I|$ , and  $(I, J) \in \mathcal{B}_0$ . We turn our attention to  $M_s$ . Applying (5.9), and the definition of the  $\Psi$ -functional in Definition 1.7, that we have

$$\sigma(I_{\theta})^{-1} \sum_{\substack{(I,J) \in \mathcal{B}_0(S) \\ J: \substack{2^s |J| = |I| \\ I_J = I_{\theta}}}} \langle H_{\sigma}(S - I_{\theta}), h_J^w \rangle_w^2 \leq \Psi^2 \psi(s)^2 \sigma(I_{\theta})^{-1} \Psi_w(F, I_{\theta}) \leq 4\psi(s)^2 \Psi^2.$$

In the last inequality, it is decisive that the interval  $I_J \subsetneq S$ , hence *fails* the condition (6.3).

**6.2. The Control of  $B_2$ .** For  $S \in \mathcal{S}$ , let  $P_S^w$  be the projection onto the span of Haar functions  $h_J^w$  with  $\pi_S J = S$ . By Lemma 5.8, there is a function  $\bar{\phi}$ , a  $w$ -Haar multiplier of  $\phi$ , so that

$$\begin{aligned} |B_2(f, \phi; S)| &\lesssim \langle H_{\sigma}(F - S), P_S^w \bar{\phi} \rangle_w \\ &= \langle H_{\sigma} F, P_S^w \bar{\phi} \rangle_w - \langle H_{\sigma} S, P_S^w \bar{\phi} \rangle_w \end{aligned}$$

Now for the first term on the right above,

$$(6.8) \quad \sum_{S \in \mathcal{S} - \{F\}} \langle H_{\sigma} F, P_S^w \bar{\phi} \rangle_w = \left\langle H_{\sigma} F, \sum_{S \in \mathcal{S} - \{F\}} P_S^w \bar{\phi} \right\rangle_w \leq H \sigma(F)^{1/2} \|P_F^w \phi\|_w$$

And, for the second term on the right above,

$$\begin{aligned} \sum_{S \in \mathcal{S}} |\langle H_{\sigma} S, P_S^w \bar{\phi} \rangle_w| &\leq H \sum_{S \in \mathcal{S}} \sigma(S)^{1/2} \|P_S^w \phi\|_w \\ &\leq H \left[ \sum_{S \in \mathcal{S}} \sigma(S) \times \sum_{S \in \mathcal{S}} \|P_S^w \phi\|_w^2 \right]^{1/2} \\ &\lesssim H \sigma(F)^{1/2} \|P_F^w \phi\|_w. \end{aligned}$$



Here, we have appealed to the critical estimate (6.4). This with (6.8) completes the bound of  $B_2(f, \phi)$ .

**6.3. The Control of  $B_1$ .** The bound for  $B_1$ , as defined in (6.6), is straightforward. Recalling our observation that the functions  $b_j$  in (6.5) are bounded in  $L^\infty$  by  $\gamma(F)$ , one can appeal directly to the testing condition to see that

$$|B_1(f, \phi; S)| \lesssim \mathbf{H} \cdot \mathbf{M} \cdot \sigma(S)^{1/2} \|P_S^w \phi\|_w,$$

$$\mathbf{M} := \sup_{J: \pi_S J = S} \left| \sum_{I: (I, J) \in \mathcal{B}(S)} \mathbb{E}_{I, J}^\sigma \Delta_I^\sigma f \right|.$$

But  $\mathbf{M} \lesssim \gamma(F)$ .

Using the orthogonality of the projections  $P_S^w$ , and the condition on the stopping intervals (6.4), one sees that

$$\begin{aligned} \sum_{S \in \mathcal{S}} |B_1(f, \phi; S)| &\lesssim \mathbf{H} \gamma(F) \left[ \sum_{S \in \mathcal{S}} \sigma(S) \times \sum_{S \in \mathcal{S}} \|P_S^w \phi\|_w^2 \right]^{1/2} \\ &\lesssim \mathbf{H} \gamma(F) \sigma(F)^{1/2} \|P_F^w \phi\|_w. \end{aligned}$$

This is as required by (6.1).

**6.4. A Second Estimate.** We have completed the proof of (6.1), and turn to the easy estimate  $\mathbf{B}\mathbf{F} \lesssim \mathbf{B}_\epsilon + \mathbf{H}$ . Indeed, if we are given a function  $f$  and  $\phi$  with which we are to test the bounded fluctuation condition, note that the sum that appears in (3.7) reduces to  $B_{\text{stop}}(f, \phi)$ . But, we have

$$\begin{aligned} |B_{\text{stop}}(f, \phi) - B_\epsilon(f, \phi)| &\leq \left| \sum_{I: I \supset F} \sum_{J \in F} \mathbb{E}_F \Delta_I^\sigma f \langle H_\sigma(I_J), g \rangle \right| \\ &\leq \left| \mathbb{E}_F f \langle H_\sigma F, g \rangle \right| + \left| \sum_{I: I \supset F} \mathbb{E}_F \Delta_I^\sigma f \langle H_\sigma(I_F - F), g \rangle \right| \end{aligned}$$

The first term is bounded by  $\mathbf{H} \sigma(F)^{1/2} \|g\|_w$ . The second term is zero, since  $f$  is supported on  $F$ , hence  $\sum_{I: I \supset F} \mathbb{E}_F \Delta_I^\sigma f \cdot (I_F - F) \equiv 0$ .

We argue that  $\mathbf{H} \lesssim \sqrt{\mathbf{A}_2} + \mathbf{W} + \mathbf{B}_\epsilon$ , which completes our proof of (3.11), that is  $\mathbf{B}\mathbf{F} \lesssim \sqrt{\mathbf{A}_2} + \mathbf{W} + \mathbf{B}_\epsilon$ .

Let us fix an interval  $I^0$ , and function  $\phi \in L^2(w)$  supported on  $I^0$ , for which we are to estimate  $\langle H_\sigma I^0, \phi \rangle_w$  in terms of the  $A_2$  constant, the weak-boundedness constant, and the split form constant  $\mathbf{B}_\epsilon$ . By appealing to the weak-boundedness constant  $\mathbf{W}$ , we can assume that  $\phi$  has  $w$ -integral zero. We are also free to consider (random) dyadic grids  $\mathcal{D}$ , with respect to which  $I^0$  is dyadic. It follows that we can take  $\phi$  in the linear span of  $\{h_J^w : J \subset I^0, J \text{ is good}\}$ .

By appealing to Theorem 3.2, it suffices to consider the sum of the two forms  $B_\ominus(I^0, \phi) + B_\epsilon(I^0, \phi)$ . But  $B_\ominus(I^0, \phi)$  is zero: If  $\Delta_J^w \phi \neq 0$ , then  $J \subset I^0$ , and so for any  $I \subset J$ , we have  $\Delta_I^\sigma I^0 = 0$ . The form  $B_\epsilon(I^0, \phi)$  is controlled by the constant  $\mathbf{B}_\epsilon$ . So our argument is complete.

## 7. DOMINATING FUNCTIONAL ENERGY

We have two estimates of the functional energy constant to prove, that  $\mathbf{F} \lesssim \mathbf{H} + \Psi$  and  $\mathbf{F} \lesssim \mathbf{H} + \mathbf{B}_\epsilon$ . The data for these considerations are a non-negative function  $f$ , its stopping intervals  $\mathcal{F}$ , and a sequence of functions  $\{g_F : F \in \mathcal{F}\}$ , with  $g_F$   $\mathcal{F}$ -adapted to  $F$ . We assume that  $\mathbf{H}$  and  $\mathbf{B}_\epsilon$  are finite, and consider the expression  $B_\epsilon(f, g_F)$ . By the monotonicity property Lemma 5.8, it suffices to assume that  $f$  takes the value  $\gamma(F)$  on each set of the form  $F - \bigcup \text{Child}_{\mathcal{F}}(F)$ . The point to observe is that if we write

$$\tilde{f}_F = f \cdot (F - \bigcup \text{Child}_{\mathcal{F}}(F)) + \sum_{F' \in \text{Child}_{\mathcal{F}}(F)} \mathbb{E}_{F'}^\sigma f \cdot F',$$

we can appeal to the bounded fluctuation property to write

$$|B_\epsilon(f \cdot F, g_F)| \lesssim \mathbf{B}\mathbf{F} \|\tilde{f}_F\|_\sigma \|g_F\|_w \lesssim \{\sqrt{\mathbf{A}_2} + \mathbf{B}_\epsilon\} \|\tilde{f}_F\|_\sigma \|g_F\|_w.$$

Applying Cauchy-Schwartz to the last two products of norms, we get

$$\begin{aligned} \sum_{F \in \mathcal{F}} \|\tilde{f}_F\|_\sigma \|g_F\|_w &\leq \left[ \sum_{F \in \mathcal{F}} \|\tilde{f}_F\|_\sigma^2 \right]^{1/2} \|g\|_w \\ &\lesssim \left[ \sum_{F \in \mathcal{F}} \sigma(F) (\mathbb{E}_F^\sigma f)^2 \right]^{1/2} \|g\|_w \lesssim \|f\|_\sigma \|g\|_w \end{aligned}$$

where we write  $g = \sum_{F \in \mathcal{F}} g_F$ . Note that the condition that the functions  $g_F$  be  $\mathcal{F}$ -adapted implies that they are orthogonal in  $F \in \mathcal{F}$ .

We conclude that on the assumption that  $\sqrt{\mathbf{A}_2}$  and  $\mathbf{B}_\epsilon$  are finite, we have

$$\sum_{F \in \mathcal{F}} |B_\epsilon(f - (f \cdot F), g_F)| \lesssim \{\sqrt{\mathbf{A}_2} + \mathbf{B}_\epsilon\} \|f\|_\sigma \|g\|_w.$$

With the assumptions on the functions  $g_F$ , in the definition of functional energy, and the inequalities (5.10) are at our disposal. This must be done in a way that controls the right-hand side of that inequality.

Let Haar multiplier  $\bar{g}_F$  of  $g_F$ , chosen so that  $\langle x, \Delta_J^w g_F \rangle_w \geq 0$  for all  $J$ . We have from (5.10) the estimate

$$\begin{aligned} (7.1) \quad \sum_{J^* \in \mathcal{J}^*(F)} P((f - (f \cdot F))\sigma, J^*) \left\langle \frac{x}{|J^*|}, g_F \right\rangle_w &\lesssim B_\epsilon(f - (f \cdot F), \bar{g}_F) + D(f, g_F) \\ D(f, g_F) &:= \sum_{t=1}^{\infty} \gamma(\pi_{\mathcal{F}}^t F) \sum_{\substack{J: J \in \mathcal{F} \\ \pi_F J = F}} \left[ \frac{|J|}{|\pi_F^t F|} \right]^{1-\epsilon} P(\sigma \pi_{\mathcal{F}}^t F, J) \sqrt{w(J)} |\langle g_F, h_J^w \rangle_w| \end{aligned}$$

Note that we are appealing to the specific form of  $f$  to obtain the form for  $D(f, g_F)$ . If the first term on the right in (7.1) is the larger, we are finished with the proof. Otherwise, we will use the  $A_2$  constant to control the terms  $D(f, g_F)$ .

Note that by repeated application of Cauchy-Schwartz in different variables, it holds that

$$\begin{aligned}
D(f, g_F)^2 &\leq \|g_F\|_w^2 \left| \sum_{t=1}^{\infty} \gamma(\pi_{\mathcal{F}}^t F) \sum_{\substack{J: J \in F \\ \pi_F J = F}} \left[ \frac{|J|}{|\pi^t F|} \right]^{1-\epsilon} P(\sigma \pi_{\mathcal{F}}^t F, J) \sqrt{w(J)} \right|^2 \\
&\lesssim \|g_F\|_w^2 \sum_{t=1}^{\infty} \gamma(\pi_{\mathcal{F}}^t F)^2 t^2 \left| \sum_{\substack{J: J \in F \\ \pi_F J = F}} \left[ \frac{|J|}{|\pi^t F|} \right]^{1-\epsilon} P(\sigma \pi_{\mathcal{F}}^t F, J) \sqrt{w(J)} \right|^2 \\
&\lesssim \|g_F\|_w^2 \sum_{t=1}^{\infty} \gamma(\pi_{\mathcal{F}}^t F)^2 t^2 \sum_{\substack{J: J \in F \\ \pi_F J = F}} \left[ \frac{|F|}{|J|} \right]^{1+\epsilon} \left[ \frac{|J|}{|\pi^t F|} \right]^{2-2\epsilon} P(\sigma \pi_{\mathcal{F}}^t F, J)^2 w(J) \\
&\lesssim A_2 \|g_F\|_w^2 \sum_{t=1}^{\infty} \gamma(\pi_{\mathcal{F}}^t F)^2 t^2 2^{-t(1-\epsilon)} \sum_{\substack{J: J \in F \\ \pi_F J = F}} \left[ \frac{|J|}{|\pi^t F|} \right]^{1-3\epsilon} P(\sigma \pi_{\mathcal{F}}^t F, J) \\
(7.2) \quad &\lesssim A_2 \|g_F\|_w^2 \sum_{t=1}^{\infty} 2^{-t/4} \gamma(\pi_{\mathcal{F}}^t F)^2 \sigma(\pi_{\mathcal{F}}^t F).
\end{aligned}$$

This holds if  $1 - 3\epsilon > \frac{1}{3}$ , which we can assume is true, as this choice of  $\epsilon$  is only associated with the definition of being good. For any  $0 < \epsilon < 1$ , we can make a choice of integer  $r$  so that it suffices to consider only  $(\epsilon, r)$ -good intervals. From the (quasi)-orthogonality conditions on  $f$  and  $\{g_F\}$ , it is then easy to see that

$$\sum_{F \in \mathcal{F}} D(f, g_F) \lesssim \sqrt{A_2} \|f\|_{\sigma} \left\| \sum_{F \in \mathcal{F}} g_F \right\|_w$$

This completes our proof of  $\mathbf{F} \lesssim \sqrt{A_2} + B_{\epsilon}$ . Indeed, to verify the last inequality, let us write  $g = \sum_{F \in \mathcal{F}} g_F$ , and apply (7.2).

$$\begin{aligned}
\sum_{F \in \mathcal{F}} D(f, g_F) &\lesssim A_2 \sum_{F \in \mathcal{F}} \|g_F\|_w \left[ \sum_{t=1}^{\infty} 2^{-t/4} \gamma(\pi_{\mathcal{F}}^t F)^2 \sigma(\pi_{\mathcal{F}}^t F) \right]^{1/2} \\
&\lesssim \sqrt{A_2} \|g\|_w \left[ \sum_{F \in \mathcal{F}} \sum_{t=1}^{\infty} 2^{-t/4} \gamma(\pi_{\mathcal{F}}^t F)^2 \sigma(\pi_{\mathcal{F}}^t F) \right]^{1/2} \\
&\lesssim \sqrt{A_2} \|g\|_w \left[ \sum_{F \in \mathcal{F}} \gamma F)^2 \sigma(F) \right]^{1/2} \lesssim \sqrt{A_2} \|f\|_{\sigma} \|g\|_w.
\end{aligned}$$

It remains to prove  $\mathbf{F} \lesssim \Psi$ , which follows from an elementary application of the side condition. We reorganize the sum around the  $\mathcal{F}$ -ancestors of an interval  $F \in \mathcal{F}$

$$\sum_{F \in \mathcal{F}} \sum_{J^* \in \mathcal{J}^*(F)} P((f(\mathbb{R} - F)\sigma, J^*) \left\langle \frac{\chi}{|J^*|}, g_F J^* \right\rangle_w$$

$$\begin{aligned}
&\leq \sum_{F \in \mathcal{F}} \mathbb{E}_F^\sigma f \cdot \sum_{t=2}^{\infty} \sum_{F' \in \mathcal{F} : \pi_{\mathcal{F}}^t = F} \sum_{J^* \in \mathcal{J}^*(F')} P(\sigma(F - F'), J^*) \left\langle \frac{x}{|J^*|}, g_{F'} J^* \right\rangle_w \\
(7.3) \quad &\lesssim \Psi \sum_{F \in \mathcal{F}} \mathbb{E}_F^\sigma f \cdot \sigma(F)^{1/2} \sum_{t=2}^{\infty} \psi(t) \left[ \sum_{F' \in \mathcal{F} : \pi_{\mathcal{F}}^t = F} \sum_{J^* \in \mathcal{J}^*(F')} \|g_{F'} \cdot J^*\|_w^2 \right]^{1/2}
\end{aligned}$$

Here, we have appealed to (1.8) with the data  $I_0 \leftarrow F$ ,  $\{I_j\} \leftarrow \{F' \in \mathcal{F} : \pi_{\mathcal{F}}^t = F\}$ , and  $\{I_{j,k}\} \leftarrow \cup\{\mathcal{J}^*(F') : F' \in \mathcal{F} : \pi_{\mathcal{F}}^t = F\}$ . It follows that we have  $\{I_{j,k}\}$  are  $(\epsilon, r)$ -good, and  $|I_{j,k}| \leq 2^{-t}|I_j|$ , and therefore, we have

$$\sum_{F' \in \mathcal{F} : \pi_{\mathcal{F}}^t = F} P(\sigma(F - F'), J^*)^2 \sum_{J^* \in \mathcal{J}^*(F')} E(w, J^*)^2 w(J^*) \leq \Psi^2 \psi(t) \sigma(F).$$

The inequality (7.3) then follows from Lemma 5.8, and the assumption that  $\mathbb{E}_{J^*}^w g_F = 0$  for all  $F \in \mathcal{F}$ , and  $J^* \in \mathcal{J}^*(F)$ . In particular, we have

$$\left| \left\langle \frac{x}{|J^*|}, g_{F'} J^* \right\rangle_w \right| \leq E(w, J^*) w(J^*)^{1/2} \|g_{F'} \cdot J^*\|_w.$$

Now, we have the quasi-orthogonality estimate (3.6). Using the condition  $\sum_{t \geq 2} \psi(t) = 1$ , we have

$$\begin{aligned}
&\sum_{F \in \mathcal{F}} \left[ \sum_{t=2}^{\infty} \psi(t) \left[ \sum_{F' \in \mathcal{F} : \pi_{\mathcal{F}}^t = F} \sum_{J^* \in \mathcal{J}^*(F')} \|g_{F'} \cdot J^*\|_w^2 \right]^{1/2} \right]^2 \\
&\leq \sum_{F \in \mathcal{F}} \sum_{t=2}^{\infty} \psi(t) \sum_{F' \in \mathcal{F} : \pi_{\mathcal{F}}^t = F} \sum_{J^* \in \mathcal{J}^*(F')} \|g_{F'} \cdot J^*\|_w^2 \leq \|g_F\|_w^2.
\end{aligned}$$

## 8. THE REMAINING ESTIMATES

**8.1. The Term  $B_{1,2}$ .** The term  $B_{1,2}$  is defined in (4.1); we are proving the estimate (4.2), where the constants on the right are the  $A_2$ -constant and the weak-boundedness constant in (3.1).

Note that by the definition of the weak-boundedness constant, we have

$$\begin{aligned}
\sum_{(I,J) : 2^{-r}|I| \leq |J| \leq 2^r|I|} \langle H_\sigma \Delta_I^\sigma f, \Delta_J^w \phi \rangle_w &\leq W \sum_{(I,J) : 2^{-r}|I| \leq |J| \leq 2^r|I|} \left| \langle f, h_I^\sigma \rangle_\sigma \langle \phi, h_J^w \rangle_w \right| \\
&\lesssim W \|f\|_\sigma \|\phi\|_w.
\end{aligned}$$

If  $3I \cap 3J = \emptyset$ , and  $|J| \lesssim |I|$ , we have the trivial consequence of (5.9),

$$(8.1) \quad |\langle H_\sigma(I), h_J^w \rangle_w| \lesssim \sigma(I) \frac{\sqrt{w(J)}|J|}{(|J| + \text{dist}(I, J))^2}$$

Now, the estimate (8.1) implies that the remaining part of  $B_{1,2}$  is controlled by

$$\sum_{(I,J) : \substack{3I \cap 3J = \emptyset \\ 2^{-r}|I| \leq |J| \leq 2^r|I|}} \langle H_\sigma \Delta_I^\sigma f, \Delta_J^w \phi \rangle_w \lesssim \sum_{(I,J) : \substack{3I \cap 3J \neq \emptyset \\ 2^{-r}|I| \leq |J| \leq 2^r|I|}} \left| \langle f, h_I^\sigma \rangle_\sigma \langle \phi, h_J^w \rangle_w \right| \cdot \alpha(I, J)$$

$$\alpha(I, J) := \frac{\sigma(I)^{1/2} w(J)^{1/2} |J|}{(|J| + \text{dist}(I, J))^2}.$$

The last coefficients satisfy the assumptions of Schur's test, with the relevant constant controlled by the  $A_2$  constant. Namely, for any  $I$  we have by the Cauchy-Schwartz inequality,

$$(8.2) \quad \left[ \sum_{J : \substack{3I \cap 3J = \emptyset \\ 2^{-r}|I| \leq |J| \leq 2^r|I|}} \alpha(I, J) \right]^2 \lesssim \sum_{(I,J) : \substack{3I \cap 3J = \emptyset \\ 2^{-r}|I| \leq |J| \leq 2^r|I|}} \frac{|J|}{(|J| + \text{dist}(I, J))^2}$$

$$\times \sum_{(I,J) : \substack{3I \cap 3J = \emptyset \\ 2^{-r}|I| \leq |J| \leq 2^r|I|}} \frac{\sigma(I) w(J)}{(|J| + \text{dist}(I, J))^2} \lesssim A_2^2.$$

This completes the proof of (4.2).

**8.2. The Term  $B_{2,1}$ .** For the term  $B_{2,1}$  of (4.3), we prove (4.5). In this sum, we have  $2^r|J| \leq |I|$ , and  $J \cap 3I = \emptyset$ . We further restrict the length of  $J$  to be  $2^{-s}|I|$ , for  $s \geq r$ . Using the estimate (8.1), we can apply the Schur test to see that

$$\sum_{(I,J) : \substack{3I \cap 3J = \emptyset \\ 2^{-r}|I| \leq |J| \leq 2^r|I|}} \langle H_\sigma \Delta_I^\sigma f, \Delta_J^w \phi \rangle_w \lesssim 2^{-s/2} A_2 \|f\|_\sigma \|\phi\|_w.$$

Indeed, the only point to observe is that for the analog of the first term on the right in (8.2), that we have the geometric decay claimed above.

**8.3. The Term  $B_{2,2}$ .** For the term  $B_{2,2}$  of (4.4), we prove the second half of (4.5). The distinction between this case and the previous is that this is the 'local' but not 'inside' part. For integers  $s \geq 1$ , we prove

$$\sum_I \sum_{J : \substack{2^s|J| = |I| \\ J \subset 3I \setminus I}} \langle H_\sigma \Delta_I^\sigma f, \Delta_J^w \phi \rangle_w \lesssim 2^{-s/2} A_2 \|f\|_\sigma \|\phi\|_w.$$

But the essential points are on the one hand that we have

$$\sum_I \sum_{J : \substack{2^s|J| = |I| \\ J \subset 3I \setminus I}} \langle \phi, h_J^w \rangle_w^2 \lesssim \|\phi\|_w^2,$$

since the length and location of  $J$  is specified by  $I$ . And on the other hand, that we have the estimate

$$\sum_{\substack{J: 2^s|J|=|I| \\ J \subset 3I \setminus I}} \langle H_\sigma h_I^\sigma, h_J^w \rangle_w^2 \lesssim A_2^2 2^{-s}.$$

To be concrete, let  $\theta, \theta' \in \{\pm\}$ , and consider the sum

$$\left| \mathbb{E}_{I_\theta}^\sigma h_I^\sigma \right|^2 \sum_{\substack{J: 2^s|J|=|I| \\ J \subset I + (\theta'|I|)}} \langle H_\sigma I_\theta, h_J^w \rangle_w^2$$

Now,  $\left| \mathbb{E}_{I_\theta}^\sigma h_I^\sigma \right| \leq \sigma(I_\theta)^{-1/2}$ , which is the estimate (2.1). And, we can apply (5.9) to see that

$$S(I) = \sum_{\substack{J: 2^s|J|=|I| \\ J \subset I + (\theta'|I|)}} \langle H_\sigma I_\theta, h_J^w \rangle_w^2 \lesssim \sum_{\substack{J: 2^s|J|=|I| \\ J \subset I + (\theta'|I|)}} P(\sigma \cdot I_\theta, J)^2 w(J)$$

We ignore the contribution from energy. But, the intervals  $J$  are good. This means that  $\text{dist}(J, I_\theta) \geq |I_\theta|^{1-\epsilon} |J|^\epsilon$ , which fact we use to estimate the Poisson integral above as follows.

**8.3. Lemma.** *Let  $J \subset I \subset I'$ ,  $|J| = 2^{-s}|I|$ , with  $s \geq r$  and  $J$  good, then*

$$P(\sigma \cdot (I' - I), J) \leq 2^{-(1-\epsilon)s} P(\sigma \cdot I', I).$$

*Proof.* Note that for  $x \in I' - I$  we have

$$\text{dist}(x, J) \geq |I|^{1-\epsilon} |J|^\epsilon = 2^{s(1-\epsilon)} |J|.$$

Using this in the definition of the Poisson integral, we get

$$\begin{aligned} P(\sigma \cdot (I' - I), J) &\leq 2 \int_{I' - I} \frac{|J|}{\text{dist}(x, J)^2} \sigma(dx) \\ &\lesssim \frac{|J|}{|I|} \int_{I' - I} \frac{|I|}{(|J| + \text{dist}(x, J))^2} \sigma(dx) \\ &\lesssim 2^{-s(1-2\epsilon)} \int_{I' - I} \frac{|I|}{(|I| + \text{dist}(x, I))^2} \sigma(dx) = 2^{-s(1-2\epsilon)} P(I, \sigma(I' - I)). \end{aligned}$$

□

Applying the Lemma, we have the estimate

$$S(I) \lesssim 2^{-2(1-2\epsilon)s} P(\sigma \cdot I_\theta, I_\theta) w(I + \theta'|I|) \lesssim A_2^2 \sigma(I_\theta).$$

**8.4. The Term  $B_{3,1}$ .** For the term  $B_{3,1}$  of (4.6), we prove (4.7). This is a simple variant of the previous estimate. Namely, we have for  $\theta \in \{\pm\}$ ,

$$\sum_{J \subset I, I_J = I_\theta} \langle H_\sigma I_\theta, h_J^w \rangle_w^2 \lesssim A_2^2 \sigma(I_\theta).$$

Here, by  $I_\theta$  we mean the child of  $I$  complementary to  $I_\theta$ . We omit the details of the argument.

## REFERENCES

- [1] Y. Belov, T. Mengestie, and K. Seip, *Discrete Hilbert transforms on sparse sequences*, available at <http://arxiv.org/abs/0912.2899>.
- [2] M. Cotlar and C. Sadosky, *On some  $L^p$  versions of the Helson-Szegő theorem*, (Chicago, Ill., 1981), Wadsworth Math. Ser., Wadsworth, Belmont, CA, 1983, pp. 306–317. MR730075 (85i:42015)
- [3] Hyun-Kyoung Kwon and Sergei Treil, *Similarity of operators and geometry of eigenvector bundles*, Publ. Mat. **53** (2009), no. 2, 417–438. MR2543858 (2010g:47011)
- [4] Michael T. Lacey, Eric T. Sawyer, and Ignacio Uriarte-Tuero, *A characterization of two weight norm inequalities for maximal singular integrals with one doubling measure*, Analysis and PDE., to appear (2008), available at <http://arxiv.org/abs/0805.0246>.
- [5] ———, *A Two Weight Inequality for the Hilbert transform Assuming an Energy Hypothesis* (2010), available at <http://www.arXiv.org/abs/1001.4043>.
- [6] Yurii I. Lyubarskii and Kristian Seip, *Complete interpolating sequences for Paley-Wiener spaces and Muckenhoupt's  $(A_p)$  condition*, Rev. Mat. Iberoamericana **13** (1997), no. 2, 361–376. MR1617649 (99e:42004)
- [7] F. Nazarov, S. Treil, and A. Volberg, *Two weight estimate for the Hilbert transform and Corona decomposition for non-doubling measures* (2004), available at <http://arxiv.org/abs/1003.1596>.
- [8] ———, *Two weight inequalities for individual Haar multipliers and other well localized operators*, Math. Res. Lett. **15** (2008), no. 3, 583–597. MR2407233 (2009e:42031)
- [9] F. Nazarov and A. Volberg, *The Bellman function, the two-weight Hilbert transform, and embeddings of the model spaces  $K_\theta$* , J. Anal. Math. **87** (2002), 385–414. Dedicated to the memory of Thomas H. Wolff.
- [10] Eric T. Sawyer, *A characterization of a two-weight norm inequality for maximal operators*, Studia Math. **75** (1982), no. 1, 1–11. MR676801 (84i:42032)
- [11] ———, *A characterization of two weight norm inequalities for fractional and Poisson integrals*, Trans. Amer. Math. Soc. **308** (1988), no. 2, 533–545. MR930072 (89d:26009)
- [12] A. Volberg, *Calderón-Zygmund capacities and operators on nonhomogeneous spaces*, CBMS Regional Conference Series in Mathematics, vol. 100, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2003.

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA GA 30332

*E-mail address:* lacey@math.gatech.edu

DEPARTMENT OF MATHEMATICS & STATISTICS, MCMASTER UNIVERSITY, 1280 MAIN STREET WEST, HAMILTON, ONTARIO, CANADA L8S 4K1

*E-mail address:* sawyer@mcmaster.ca

DEPARTMENT OF MATHEMATICS & STATISTICS, MCMASTER UNIVERSITY, 1280 MAIN STREET WEST, HAMILTON, ONTARIO, CANADA L8S 4K1

*E-mail address:* cyshen@mcmaster.ca

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, EAST LANSING MI

*E-mail address:* ignacio@math.msu.edu